DOCUMENT RESUME

ED 334 196 TM 016 046

AUTHOR Yusuf, Mian Muhammad; And Others
TITLE Characteristics of Multiple Comparison Procedures.
PUB DATE Oct 87
NOTE 23p.; Paper presented at the Annual Meeting of the Midwestern Educational Research Association (Chicago, IL, October 15-17, 1987).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Analysis of Variance; *Comparative Analysis; Robustness (Statistics); Tables (Data)
IDENTIFIERS *Multiple Comparison Tests

ABSTRACT

Multiple comparison procedures (MCPs), in the context of analysis of variance, are becoming more popular than the overall F tests. An attempt is made to clarify confusion among the different MCPs by systematically comparing and contrasting the procedures in terms of their purposes, restrictions, robustness to assumptions, and other special features that are relevant to any analysis of variance design. Focus is on providing a flowchart to assist researchers in making sound decisions regarding the choice of a MCP. The MCPs compared include: (1) multiple t; (2) multiple F; (3) Dunn; (4) Dunn-Sidak; (5) Dunnett; (6) Fisher's LSD; (7) Tukey's HSD; (8) Spjotvoll-Stoline's t; (9) Tukey-Kramer; (10) Games-Howell; (11) Tamhane; (12) Scheffe's; (13) Brown-Forsythe; (14) Newman-Keuls; and (15) Duncan. Table 1 summarizes characteristics of these procedures, and Table 2 presents a flowchart. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED334196

CHARACTERISTICS OF MULTIPLE COMPARISON PROCEDURES

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

By

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

MIAN MUHAMMAD YUSUF

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Mian Muhammad Yusuf
University of Wisconsin-Parkside, Kenosha
WI 53141-2000

Chao-Ying J. Peng
Indiana University, Bloomington
IN 47401

Yuung-Shi Shyuu
Indiana University, Bloomington
IN 47401

Paper presented at the annual conference of the
Mid-Western Educational Research Association
Chicago, Illinois October 16, 1987

TM016040

BEST COPY AVAILABLE

ABSTRACT

Multiple comparison procedures, in the context of analysis of variance, are becoming more popular than the overall F tests. It is due in part to the power of these comparison procedures as well as their control of the Type I error rate. Unfortunately these procedures are not equally understandable. This paper attempts to clarify the confusion by systematically comparing and contrasting these procedures in the terms of their purposes, restrictions, robustness to assumptions and other special features. Finally a flow chart is presented to help readers make a proper choice regarding comparison procedures applied to the comparison among the treatment means.

CHARACTERISTICS OF MULTIPLE COMPARISON PROCEDURES

OBJECTIVES:

The specific objectives of this paper are:

1. To compare and contrast multiple comparisons procedures on characteristics that are relevant to any analysis of variance (ANOVA) design.

2. To provide a flow chart which will assist researchers in making sound decisions regarding the choice of a multiple comparison procedure.

INTRODUCTION:

Since 1950's the topic of multiple comparison procedures has gained increasing popularity among behavioral data analysts. The trend is partly due to their control of Type 1 error rate. Along with this popularity came the problem of confusion as to **when** to use **which** of these comparison procedures. Several review articles in the past (Harter, 1957; Ryan, 1972; Tukey, 1953) concentrate on only a few characteristics of these procedures. A more recent reference (Kirk, 1982) improves upon previous articles by including more procedures to be contrasted and by clearly listing the aspects

on which these procedures differ (pp. 128 and 129). Yet this reference fails to inform users on how to arrive at a proper choice of a multiple comparison procedure within the constraints of their research design.

The primary objective of this paper , therefore, is to expand Kirk's table to include information about other aspects of the multiple comparison procedures. Also, a flow chart is presented in this paper to assist users in reaching a proper choice on multiple comparison procedures.

The procedures chosen for discussion include Multiple t, Multiple F, the Dunn, the Dunn-Sidak, the Dunnett, the Fisher's LSD, the Tukey's HSD, the Spjøtvoll-Stoline's t, the Tukey-Kramer, the Games-Howell, the Tamhane, the Scheffe's, the Brown-Forsythe, the Newman-Keuls, and the Duncan procedures. The characteristics by which these procedures are contrasted are purpose (i.e., simple vs. complex contrast), the requirement of equal n per group, homogeneous vs. heterogeneous variance, the sampling distribution and unique features (see Table 1).

Finally , a flow chart is provided in Table 2. The flow chart is a system of decision-making processes which ultimately lead users to one or two procedures which are recommended based on the information revealed in Table 1.

Table 1

SUMMARY OF CHARACTERISTICS OF MULTIPLE COMPARISON PROCEDURES

1	2	3	4	5	6	7	8
PROCEDURE	**SIM. or COM.**	**EQU. n's or UNEQU. n's**	**HOM. or HET.**	**USE TABLE**	**CRITICAL VALUE**	**SPECIAL FEATURE**	**FORMULA**
A: PRIORI ORTHOGONAL CONTRASTS							
1. MULTIPLE t	**COM.**	**UNEQU.**	**HOM.**	**KIRK E.4 p813**	**Student's t**	**For planned Orthogonal Contrasts**	$t = \frac{\Psi_i}{\hat{\sigma}_{\Psi_i}} = \frac{\sum_{j=1}^{p} c_j \bar{Y}_{\cdot j}}{\sqrt{MS_e \sum_{j=1}^{p} \frac{c_j^2}{n_j}}}$

1	2	3	4	5	6	7	8
PROCEDURE	**SIM. or COM.**	**EQU. n's or UNEQU. n's**	**HOM. or HET.**	**USE TABLE**	**CRITICAL VALUE**	**SPECIAL FEATURE**	**FORMULA**
2. **MULTIPLE t**	**COM.**	**UNEQU.**	**HOM.**	**KIRK E.5 p814**	**F Distribution**	**V=1 V_2=p(n-1) for equal n's V_2= N-p for unequal n's**	$F=\frac{\Psi_i^2}{\hat{\sigma}^2_{\Psi_i}}$ $=\frac{\left(\sum_{j=1}^{p} C_j \bar{Y}_{\cdot j}\right)^2}{MS_e \sum_{j=1}^{p} \frac{C_j^2}{n_j}}$
3. **MULTIPLE t** **With Behrens Fisher and Welch Procedures**	**COM.**	**UNEQU.**	**HET.**	**ASPIN t′**	**Student's t/t′ distribution**	**Good against type I error when the variances are hetrogeneous**	$t' = \frac{\Psi_i}{\hat{\sigma}_{\Psi_i}}$ $= \frac{C_j \bar{Y}_{\cdot j} + C_{j'} \bar{Y}_{\cdot j'}}{\sqrt{\frac{\hat{\sigma}_j^2}{n_j} + \frac{\hat{\sigma}_{j'}^2}{n_{j'}}}}$

1	2	3	4	5	6	7	8
PROCEDURE	SIM. or COM.	EQU. n's or UNEQU. n's	HOM. or HET.	USE TABLE	CRITICAL VALUE	SPECIAL FEATURE	FORMULA
B: PRIORI NONORTHOGONAL CONTRASTS							
4. DUNN D	COM.	UNEQU.	HOM.	KIRK E.16 p842	Percentage Points of DUNN MULTIPLE COMPARISON TEST	Useful for C planned contrast.	tD = t as in 1
5. DUNN-SIDAK DS	COM.	UNEQU.	HOM.	KIRK E.17 p843	Percentage Points of DUNN-SIDAK MULTIPLE COMPARISON TEST	Slightly powerful than DUNN	tDS = t as in 1

10

11

1	2	3	4	5	6	7	8
PROCEDURE	**SIM. or COM.**	**EQU. n's or UNEQU. n's**	**HOM. or HET.**	**USE TABLE**	**CRITICAL VALUE**	**SPECIAL FEATURE**	**FORMULA**
6. **DUNNET tD′**	**SIM.**	**UNEQU. & EQU.**	**HOM. HET***	**KIRK E.9 p826**	**Percentage Points for the comparison of p-1 treatment means with a control**	**Slightly powerful than DUNN** **<u>Two condition apply:</u>** **1) p-1(GE-GC) means compared with a Control Group (GC)** **2) Correlation between any two contrast = 0.5**	$tD' = \frac{\Psi_i}{\partial \Psi_i}$ $= \frac{C_j \bar{Y}_{.j} + C_{j'} \bar{Y}_{.j'}}{\sqrt{MS_e \left(\frac{C_j^2}{n_j} + \frac{C_{j'}^2}{n_{j'}} \right)}}$

1	2	3	4	5	6	7	8
PROCEDURE	**SIM. or COM.**	**EQU. n's or UNEQU. n's**	**HOM. or HET.**	**USE TABLE**	**CRITICAL VALUE**	**SPECIAL FEATURE**	**FORMULA**
C: A POSTERIORI NONORTHOGONAL CONTRASTS							
7. FISHER LSD	**SIM.**	**UNEQU.**	**HOM.**	**KIRK p813**	**Student's t**	**Appropriate with the restriction that the ANOVA F value be significant at $\alpha=0.05$**	$\Psi(LSD) = t_{\alpha/2,\nu}\sqrt{\frac{2MS_e}{n}}$
8. TUKEY HSD	**SIM.**	**EQU.**	**HOM.**	**KIRK p822**	**$q/\sqrt{2}$**	**Best for the largest pairwise diff. between means an alternative to the overall F test.**	$\Psi(HSD) = q_{\alpha;p,\nu}\sqrt{\frac{MS_e}{n}}$
9. SPJØTVOLL-STOLINE T'	**SIM.**	**UNEQU.**	**HOM.**	**KIRK p846**	**$T'/2$**	**The n's are similar**	$\Psi(T') = q'_{\alpha;p,\nu}\sqrt{\frac{MS_e}{n_{min}}}$

1	2	3	4	5	6	7	8
PROCEDURE	**SIM. or COM.**	**EQU. n's or UNEQU. n's**	**HOM. or HET.**	**USE TABLE**	**CRITICAL VALUE**	**SPECIAL FEATURE**	**FORMULA**
10. **TUKEY-KRAMER TK**	**SIM.**	**UNEQU.**	**HOM.**	**KIRK p822**	**Studentized Range**	**Recommended for a moderate or large imbalance among the sample n's**	$\Psi(TK) = q_{\alpha;p,\nu}\sqrt{MS_e\left(\frac{1}{n_j}+\frac{1}{n_{j'}}\right)\Big/2}$
11. **GAMES-HOWELL GH**	**SIM.**	**UNEQU.**	**HET.**	**KIRK p822**	**Studentized Range**	**Slightly greater power than Tamhane**	$\Psi(GH) = q_{\alpha;p,\nu'}\sqrt{\left(\frac{\hat{\sigma}_j^2}{n_j}+\frac{\hat{\sigma}_{j'}^2}{n_{j'}}\right)\Big/2}$

1	2	3	4	5	6	7	8
PROCEDURE	**SIM. or COM.**	**EQU. n's or UNEQU. n's**	**HOM. or HET.**	**USE TABLE**	**CRITICAL VALUE**	**SPECIAL FEATURE**	**FORMULA**
12. **TAMHANE T2**	**SIM.**	**UNEQU.**	**HET.**	**KIRK p843**	**tD**	**Better control of Exp. wise error rate ($\mathcal{L}_{EW}$) then GH**	$\Psi(T2) = tDS_{\alpha/2;c,v'} \sqrt{\frac{\theta_j^2}{n_j} + \frac{\theta_{j'}^2}{n_{j'}}}$
13. **SCHEFEE'S S**	**COM.**	**UNEQU.**	**HOM.**	**KIRK p814**	**S**	**Advantage to evaluate all possible contrasts.**	$\Psi(S) = \sqrt{V_1 F_{\alpha;v_1,v_2}} \sqrt{MS_e \sum_{j=1}^{p} \frac{C_j^2}{n_j}}$ $V_1 = (p-1)$
14. **BROWN-FORSYTHE BF**	**COM.**	**UNEQU.**	**HET.**	**KIRK p814**	**F**	**A modification of Scheffe's S, but more robust with heterogeneity of variance.**	$\Psi(BF) = \sqrt{V_1 F_{\alpha;v_1,v'_2}} \sqrt{\sum_{j=1}^{p} \frac{C_j^2 \theta_j^2}{n_j}}$ $V_1 = (p-1)$

1	2	3	4	5	6	7	8
PROCEDURE	SIM. or COM.	EQU. n's or UNEQU. n's	HOM. or HET.	USE TABLE	CRITICAL VALUE	SPECIAL FEATURE	FORMULA
15. NEWMAN-KEULS (NK)	SIM.	EQU.	HOM.	KIRK p822	$q / \sqrt{2}$	Based on a stepwise or layer approach to significance testing. More powerful than Tukey's Test.	$\Phi(W_r) = q_{\alpha;r,\nu}\sqrt{\frac{MS_e}{n}}$
16. DUNCAN	SIM.	EQU.	HOM.	KIRK p824	$t_n / \sqrt{2}$	Share many features of the NK including the stepwise approach to significance testing, but more powerful as p (number of group) increases.	$\Phi(W_r) = q_{\alpha;r,\nu}\sqrt{\frac{MS_e}{n}}$

NB 1. Number of Pairwise (simple contrast) = p(p-1)/2 (p= # of Group Means)
2. Number of ORTHOGONAL CONTRAST = p-1 for a set.

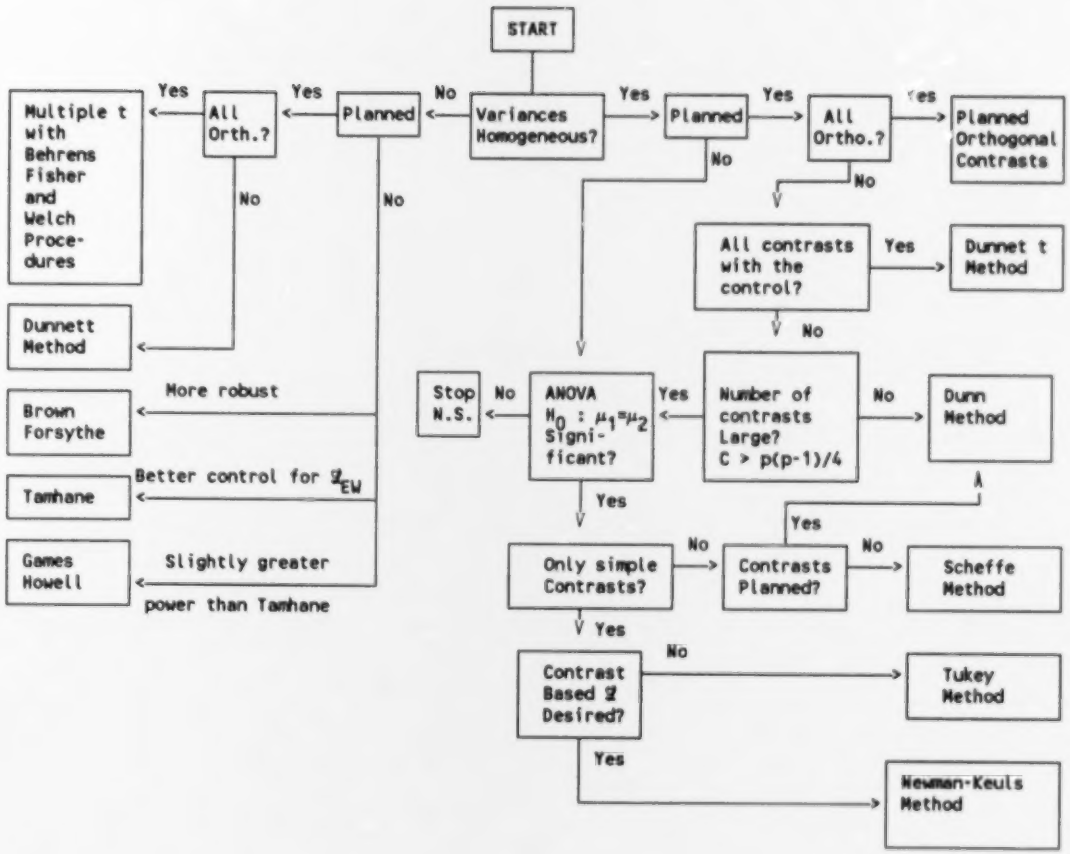
• With Modification.

ABBREVIATIONS:

UNEQU. = UNEQUAL n's HET. = HETEROGENEOUS VARIANCES SIM. = SIMPLE COMPARISONS
EQU. = EQUAL n's HOM. = HOMOGENEOUS VARIANCES COM. = COMPLEX COMPARISONS

Table 2

Flow Chart



References

Fisher, R. A. (1949). The design of experiments. Edinburgh: Oliver and Boyd Ltd.

Glass, G. V., and Hopkins, K. D. (1984). Statistical Methods in Education and Psychology. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Harter, H. L. (1957). Error rates and sample sizes for range tests in multiple comparisons. Biometrics, 13, 511-536.

Kirk, R. E. (1982). Experimental design: Procedures for the behavioral sciences. Belmont, CA: Brooks/Cole Publishing Company.

Ryan, T. A. (1972). Multiple comparisons in psychological research. In R. E. Kirk (ed.), Statistical issues: A reader for the behavioral sciences. Montersy, CA: Brooks/Cole Publishing Company.

Tukey, J. W. (1953). The problem of multiple comparison. Privately circulated monograph.